EARLY EVOLUTION OF CELLULAR ELECTRON TRANSPORT: MOLECULAR MODELS FOR THE FERREDOXIN-RUBREDOXIN-FLAVODOXIN REGION

GUNNAR VON HEIJNE* and CLAS BLOMBERG
Research Group for Theoretical Biophysics, Dept. of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

and

HERRICK BALTSCHEFFSKY
Dept. of Biochemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm, Sweden

(Received 21 November, 1977)

Abstract. Using information from conformation-predicting algorithms and X-ray data, a possible mode of structural evolution from an ancestral molecule of about 28 residues to bacterial ferredoxins and rubredoxins is suggested. The possibility of a further evolutionary pathway leading to flavodoxin-like proteins is also indicated.

1. Introduction

Bacterial ferredoxins and rubredoxins belong to a large group of iron-sulphur proteins (Hall *et al.*, 1974). The smallest of these electron-transporting proteins are 52–55 residues long. Ferredoxin binds one or two 4Fe + 4S clusters per molecule, rubredoxin only one Fe atom. Functionally they are redox-proteins, ferredoxins with a redox-potential of around –400 mV (close to that of the hydrogen electrode), the value for rubredoxins being much higher, often around –60 mV. It has been suggested that there should exist an evolutionary link between ferredoxin and rubredoxin, in agreement with a hypothesis assuming stepwise evolution along and across the potential scale (Baltscheffsky, 1974a, b). This suggestion has recently found support from new amino acid sequence data indicating such a link between the ferredoxins and the rubredoxins (Vogel *et al.*, 1978; Yčas, 1976). Since evolution works on the three-dimensional structure of a protein even more than on its amino acid sequence, this evolutionary connection should be understandable in terms of a model for the three-dimensional structural relationships between the two proteins.

It is the main purpose of this article to show how the concept of temporary structures introduced in Section 2 and more fully described elsewhere (in preparation) makes it possible to propose a simple model for a structural evolution from a 28-residue ancestral unit to ferredoxin and rubredoxin. In fact, such a short ferredoxin half-unit has been demonstrated to have some, although lowered, electron transporting capacity (Orme-

* To whom reprint requests should be addressed.

Origins of Life 9 (1978) 27–37. *All Rights Reserved.*
Copyright © 1978 by *D. Reidel Publishing Company, Dordrecht, Holland.*

Johnson, 1972). The possibility of a further evolutionary pathway leading to a flavodoxin-like protein (Baltscheffsky, loc. cit.) is also briefly discussed.

In this context we wish to stress the importance of the anti-parallel β-pleated sheet as a good candidate for an early stable element of polypeptide structure to occur on the primitive earth (Orgel, 1972; Carter and Kraut, 1974; Brack and Orgel, 1975; Richardson, 1977). In particular, this type of structure, which is known to exist in rubredoxin, will be postulated for the ferredoxin apoprotein.

2. Evolution of 'Primitive' Proteins

An organized protein structure is the result of a more or less well-defined folding process (Anfinsen and Scheraga, 1975). Thus, the evolution of structure should be intimately related to the evolution of folding pathways. Such pathways are generally described as hierarchic processes starting from short segments of the protein chain with a slight inherent tendency to adopt certain local conformations, e.g. helix, extended or bend (Anfinsen and Scheraga, loc. cit.; Karplus and Weaver, 1976; Robson and Pain, 1976; Rose *et al.*, 1976). In the succeeding step these conformations may become somewhat stabilized as a result of pairwise interactions between the segments, yielding semi-local elements of tertiary structure such as sections of β-sheet or interacting helices. Finally, these structures together with remaining coil-parts of the chain 'condense' into the native structure, stabilized by, for example, specific hydrogen bonds and hydrophobic interactions (Chothia, 1976).

However, the situation is different for short polypeptides (much less than 100 residues) or for those which are poorly evolved in the sense that they lack specific stabilizing hydrogen bonds. Then, weak interactions of the kind mentioned above will probably not be enough to stabilize the functional structure, and some covalent stabilization will be needed, e.g. sulphur bridges. The rate of formation of such stabilizing bonds at particular places will be strongly enhanced by the existence of semi-local structural elements with the reacting groups in the proper orientation. As will be discussed in a forthcoming paper, the life time of this type of structure will be long enough to facilitate the formation of these bonds but not so long as to allow biological function. Such structures, which should be detectable experimentally, will be called temporary structures.

As a simple illustration, let us consider the formation of a two-stranded anti-parallel β-pleated sheet in a molecule some 40 residues long. This structure may be described by two equilibrium constants: that for bend formation, called K_i, and that for the individual intra-strand residue-residue interactions, K_p. Representative values are $K_i \approx 2 \cdot 10^{-2}$ and $K_p \approx 1.2–1.3$ (Lewis *et al.*, 1973; Birshtein *et al.*, 1976). For a sheet with 20 residue-residue pairs this gives a total equilibrium constant for the coil⇔sheet equilibrium equal to $K = K_i \cdot K_p{}^{20} \approx 0.8–3.8$ which seem to be quite reasonable values considering the proposed functional role of such a structure.

In principle, for 'primitive' polypeptides, the local structural elements in a given

molecule may well have been able to interact in more than one way, thus establishing an equilibrium distribution of several different temporary structures. If stabilizing covalent bonds could be formed efficiently in more than one of these structures, the final product would consist of two or more different stable structures in amounts reflecting the details of the process. These amounts would be influenced by mutational events affecting the folding, and the most useful structure (or structures) would eventually be selected for. Finally, after a gene duplication, a minor but potentially useful component of this equilibrium distribution could be selected for in one of the resulting genes, giving rise to a protein resembling the original one only at the level of local structural elements. Recently, the idea that new function may arise prior to a gene duplication was also proposed by Orgel (1977).

In the next section this kind of reasoning will be applied in an attempt to elucidate the apparent evolutionary relationships between ferredoxin and rubredoxin on the basis of current knowledge about the structures of the pertinent proteins.

3. Ferredoxin – Rubredoxin Evolution

When looking for structural similarities among small proteins such as the ferredoxins and rubredoxins, our apparently novel approach has been that such similarities should be sought for preferentially in the apoproteins since the *holo*-forms may be more or less distorted by the binding of functional groups, possibly influencing a considerable proportion of the residues. As has been argued in Section 2, small apoproteins may spend a substantial fraction of their time in one or a few temporary structures built up from local structural elements such as stretches of residues in extended, helical or bend conformations. In the case of apo-ferredoxin the existence of at least one temporary structure can be inferred from the following observations: (i) H^3-labeled apoferredoxin quickly exchanges all its hydrogens under H^3-H^1 exchange conditions (Hong and Rabinowitz, 1970) indicating the absence of any long-lived structure; (ii) active ferredoxin can be reconstituted from apo-ferredoxin in high yield, 70–90% (Hong and Rabinowitz, 1967); and (iii) the two Fe-S clusters bind to cysteine residues that are far apart in the sequence; the specific binding pattern revealed in the X-ray structure (Adman *et al.*, 1973, 1975) clearly shows that the clusters do not bind to a random coil protein.

With this in mind, eight bacterial ferredoxins were analyzed for local structural elements by two prediction methods (Burgess *et al.*, 1974; Chou and Fasman, 1974), an approach intended to increase the reliability of the structure prediction that has also been used by other workers (Gabel *et al.*, 1976), cf. Lenstra *et al.* (1977). The results are presented in Figure 1. As can be seen from the figure, a large fraction of the residues are predicted to be in an extended conformation. Due to the difficulty to make a definite assignment, better reliability was sought, however.

The accuracy of the data can indeed be improved if account is taken of the facts that ferredoxins have an approximately twofold symmetry in the X-ray holoprotein structure

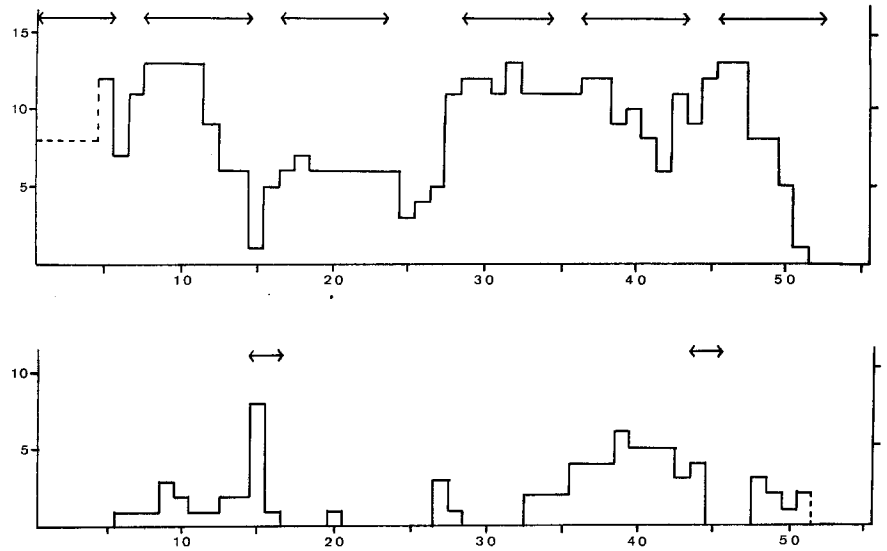

Fig. 1a–b. Joint secondary structure prediction for eight bacterial ferredoxins: *Peptococcus aerogenes, Clostridium pasteurianum, Clostridium acidiurici, Peptostreptococcus elsdenii, Clostridium tartarivorum, Clostridium thermosaccharolyticum* (Perutz and Raidt, 1975), *Clostridium butyricum* and *Clostridium M–E* (Tanaka *et al.*, 1974). Since two prediction methods were used (Burgess *et al.*, 1974; Chou and Fasman, 1974) each position was analyzed sixteen times. In figure (a) the number of times that each residue position was predicted to be in an extended conformation is plotted against the residue number, and the bend predictions are shown in (b) in the same manner (only the method of Burgess *et al.* was used to predict bends). Helix predictions are not shown since only three of the analyzed proteins were predicted to have helical sections.

(Adman *et al.*, loc. cit.) and that the amino acid sequences clearly indicate an earlier gene duplication (Tanaka *et al.*, 1971). Thus, if the two halves in Figure 1 are added (with due account taken of an insertion between residues 32 and 33 and a deletion of residues 55 and 56) Figure 2 is obtained, and it is possible to identify even more clearly the indicated local structural elements in the 28-residue unit, Table I. Duplication of this unit then leads to a prediction for the apoferredoxin, Table I and Figure 1. These predicted structures have the general characteristics expected for β-sheet structures (Sternberg, Thornton, 1977).

Considering now the binding between prosthetic group and *apo*-protein, the two 4Fe + 4S clusters bind to cysteine residues 8–11–14–47 and 18–37–40–43, respectively. This makes it possible to propose 'active', cluster-binding temporary structures for the apoferredoxin as well as for the *apo*-form of the ferredoxin half-unit, Figure 3a, b.

As far as the cluster binding process is concerned it is to be noted that the distance between the two cysteine side-chains represented by unfilled circles in Figure 3b is approximately 8.5 Å (Ananthanarayanan, 1972), allowing the distance between the two sulphur atoms to vary from 5.1 Å to 11.8 Å. When bonded to the cluster this distance is

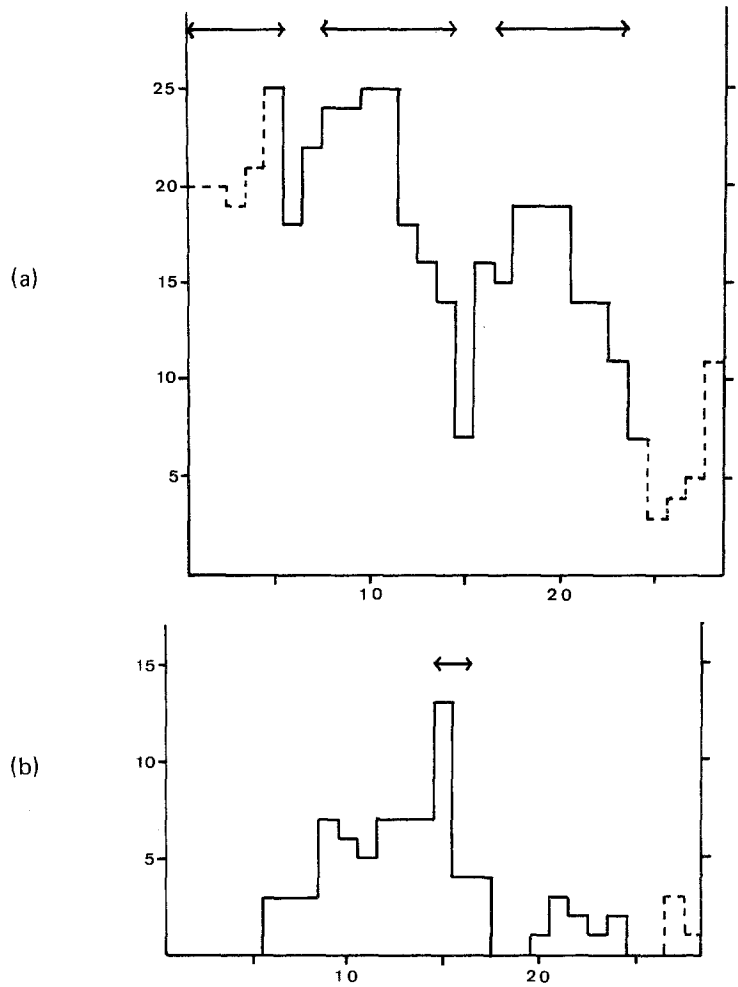

Fig. 2a–b. Joint secondary structure prediction for the 28-residue ferredoxin half-unit (see text). (a) extended structure; (b) bends.

~8.3 Å; these two residues have an almost ideal spacing for binding to an incoming cluster. We therefore propose that the clusters first bind to residues 11,47 and 18,40 respectively, whereupon the sheet is ruptured and the clusters are free to rotate ~180° down between the strands to positions where they can form the remaining bonds. If cysteines 11,18,40 and 47 are taken to be projecting up from the plane of the paper in Figure 3b, this binding process leads to the correct stereo-chemical cluster-binding pattern, as compared with the known three-dimensional structure of *Peptococcus aerogenes* ferredoxin. Incidentally, this kind of argumentation has earlier been used in a description of a possible folding mechanism for the oxidized *Chromatium* high potential iron protein (Carter *et al.*, 1974).

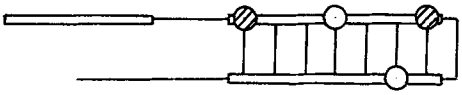

Fig. 3a.

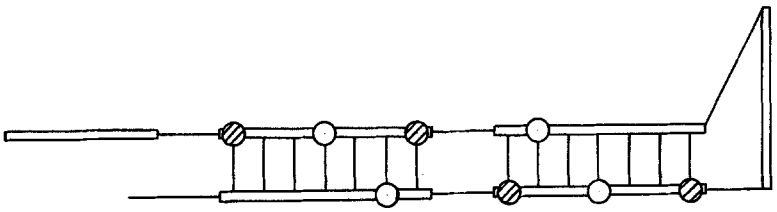

Fig. 3b.

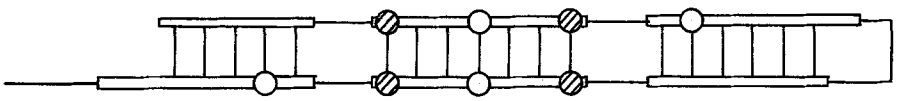

Fig. 3c.

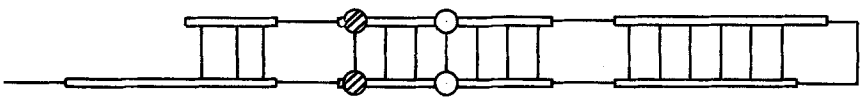

Fig. 3d.

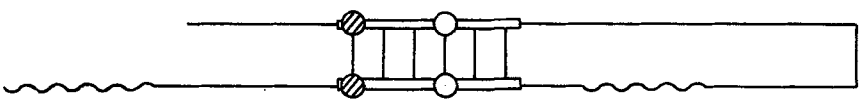

Fig. 3e.

Fig. 3a–e. Proposed temporary apoprotein structures. Stretches of extended structure are represented by heavy bars, helices by wavy lines and hydrogen-bonds by thin lines. Filled and unfilled circles represent cys-residues with side-groups projecting out from different sides of the sheet. The amino-end of the protein is always at the upper left-hand corner. (a) ferredoxin half-unit; (b) ferredoxin; (c) alternative, rubredoxin-like structure; (d) rubredoxin, based on (c); (e) rubredoxin, based on secondary structure predictions for contemporary rubredoxins.

The proposed apoferredoxin structure in Figure 3b may be compared with the known three-dimensional structure of *P. aerogenes* ferredoxin (Adman *et al.*, loc. cit.). The predicted sections of extended structure 1–5, 17–20 and 29–34 are readily discernible, and the two cluster binding sections 8–14 and 37–43 have a slightly bent extended conformation with side-chains protruding in alternate directions (Adman *et al.*, 1973) as is to be expected if they were in a pleated sheet in the apoprotein. The proposed apoferredoxin structure in Figure 3b thus essentially resembles the actual holoprotein structure.

TABLE I

Predicted local structural elements for bacterial ferredoxins and rubredoxins

	Helix	Extended	Bend
Ferredoxin half-unit[a]	–	1–5, 8–14, 17–23	15–16
Ferredoxin half-unit[b]	–	1–5, 8–12, 16–23	13–15, 24
Ferredoxin	–	1–5, 8–14, 17–23, 29–34, 37–43, 46–52	15–16, 44–45
Rubredoxin[a]	–	1–3, 6–12, 15–21 27–32, 35–41, 44–50	13–14, 42–43
Rubredoxin[c]	30–33, 48–52	6–11, 36–41	20–22, 25–27 34–35, 46–47

[a] Predicted on the basis of results for ferredoxin (see text).
[b] Predicted from an ancestral ferredoxin half-unit sequence adapted from Yčas (1976) (see text).
[c] Predicted from contemporary rubredoxin amino-acid sequences.

In a recent article (Yčas, 1976), an ancestral sequence for the bacterial ferredoxin half-unit was proposed. If residues 2,6,10 and 14–22, which we consider to be probable insertions, are excluded from this sequence a 28-residue half-unit similar to the one discussed above is obtained. Using the prediction methods described, a remarkably good agreement with the earlier half-unit prediction is found, Table I.

An alternative temporary structure is shown in Figure 3c. The cysteine residues in the middle sheet section are now closer together, the distance between the side-chains 8 and 43 (11 and 40) being ~4.8 Å (Ananthanarayanan, loc. cit.) allowing the S–S distance to vary between 1.5 Å and 8.1 Å. This short distance makes it difficult to bind a 4Fe + 4S cluster but it facilitates the binding of a single iron atom (as in rubredoxin) where the appropriate S–S distance is ~3.8 Å.

We therefore suggest that four of the cysteine residues in Figure 3c could bind an iron atom whereas the remaining lost their functional significance and were discarded in the further evolutionary development. Indeed, contemporary rubredoxins generally have about 52 residues with cysteine in positions 6,9,38 and 41, a situation which arises if two residues in Section 1–7 and one in Section 48–55 are deleted in Figure 3c, giving the temporary structure for aporubredoxin shown in Figure 3d. This structure may be compared with the results from a search for local structural elements in seven rubredoxins based on the same method as above, Figure 4, leading to the predicted temporary structure in Figure 3e which reproduces essential features of the published X-ray structure for *Clostridium pasteurianum* rubredoxin (Herriott *et al.*, 1970).

Thus we propose the simple model shown in Figure 3 for the structural evolution from a small, 28-residue half-unit leading to ferredoxin and rubredoxin based on the temporary structures presented.

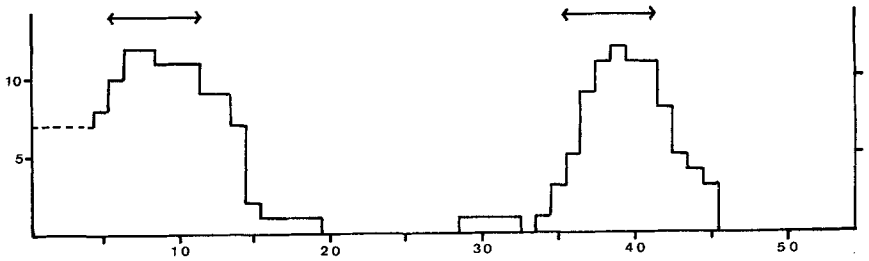

Fig. 4a.

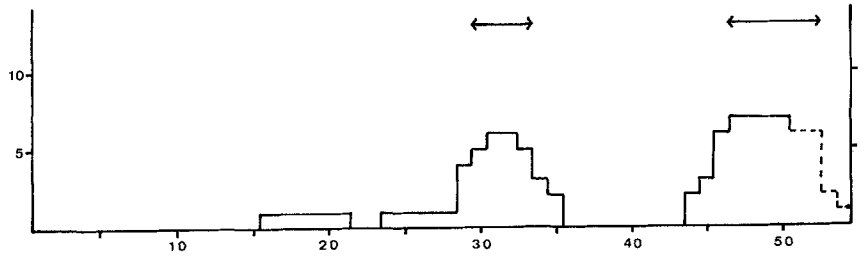

Fig. 4b.

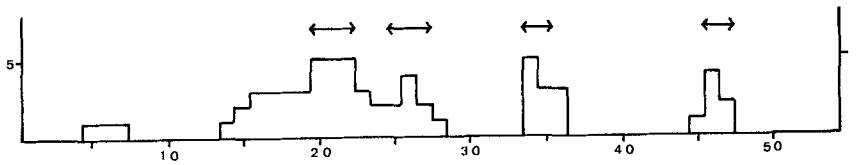

Fig. 4c.

Fig. 4a–c. Joint secondary structure prediction for six bacterial rubredoxins: *Clostridium pasteurianum* (McCarthy, 1972), *Desulfovibrio gigas* (Bruschi, 1976a), *Desulfovibrio vulgaris* (Bruschi, 1976b), *Pseudomonas oleovorans*, res. 1.54 (Benson *et al.*, 1971), *Micrococcus aerogenes* (Bachmayer *et al.*, 1968a) and *Peptostreptococcus elsdenii* (Bachmayer *et al.*, 1968b). (a) extended structure; (b) helical structure; (c) bends. See legend, Figure 1, for further details.

4. Possible Structural Connections Between Rubredoxin and Flavodoxin

As was pointed out by one of us (Baltscheffsky, loc. cit.), evolution across the potential scale may have led from iron-binding proteins to nucleotide-binding proteins. Some indications that rubredoxin and flavodoxin may be evolutionary related were given. Flavodoxin has a parallel pleated sheet structure (Andersen *et al.*, 1972), and thus one may ask whether, possibly, an anti-parallel pleated sheet structure, similar to the one seen in rubredoxin, may have evolved into a corresponding parallel pleated sheet arrangement in flavodoxin.

Indeed, the aporubredoxin structure in Figure 3e indicates the possibility of another temporary structure for this protein (or at least for a less developed ancestor), Figure 5. It may parenthetically be noted that the loop sequence connecting the two sheet strands in

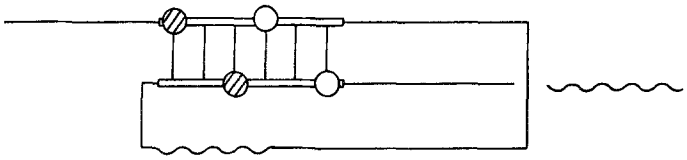

Fig. 5. Alternative temporary structure for rubredoxin with a parallel sheet, cf. Figure 3e.

this structure is 24 residues long whereas the connecting sequences in flavodoxin are 22 or 23 residues long, and that the sheet strands are of comparable lengths: 6 residues in rubredoxin, 6 or 7 in flavodoxin (Andersen *et al.*, loc. cit.). Moreover, extension of the short helical Section 30–33 in rubredoxin towards the C-terminus could possibly give rise to a β-α-β fold, typical for the nucleotide-binding proteins. This may be a first indication of a three-dimensional structural relationship between the iron-containing and the nucleotide-binding proteins. However, a quite different evolutionary origin for flavodoxin has recently been suggested by Fox (1976) and Kobayashi and Fox (1977). Clearly, more work is needed before this difficult question can be settled.

5. Discussion

The concept of temporary structures appears to be useful in the search for structural evolution among small proteins. Specifically, it has made it possible to propose a detailed model for the evolution from a small ferredoxin half-unit to ferredoxin and rubredoxin. This model accounts well for the known *holo*-enzyme structures, and the Fe-binding cysteine residues in the proposed *apo*-enzyme structures turn out to have an almost optimal geometry for ease of binding to the prosthetic group. Moreover, the proposed evolutionary pathway does not involve any improbable structural alterations.

Also, the possibility that an evolutionary link exists between rubredoxin and flavodoxin has been discussed in some detail.

A novel result of the present work is the finding with ferredoxins that a predicted β-structure forming capability by a polypeptide chain may be unrealized in the final configuration of a protein due to its binding of structures such as the distorted cube iron-sulfur clusters of the bacterial ferredoxins. It should be possible to test experimentally the question of whether the apoproteins indeed contain the predicted β-structures, with already existing methods for reversible removal of the iron and sulfur from ferredoxins (Malkin and Rabinowitz, 1966; Thompson, C. L. *et al.*, 1974). With ferredoxin apoproteins obtained in this way, circular dichroism, and X-ray data if suitable apoprotein crystals can be made, would provide the necessary information. We hope to tackle this problem in the near future.

If the predicted antiparallel β-structure pattern indeed exists in the apoprotein of ferredoxin, then one is directly led to the more general question: how may binding of various groups by apoproteins change secondary structural properties in the functional

holoprotein? At the present time, even before direct experimental evidence has been obtained, the results from our theoretical analysis would appear to underline that the possibilities for such changes are important to consider.

References

Adman, E. T., Sieker, L. C., and Jensen, L. H.: 1973, *J. Biol. Chem.* **248**, 3987–3996.

Adman, E. T., Sieker, L. C., and Jensen, L. H.: 1975, 'in '10th International Congress of Crystallography Abstracts', p. 34, Amsterdam.

Ananthanarayanan, V. S.: 1972, *J. Scient. Ind. Res.* **31**, 593–612.

Andersen, R. D., Apgar, P. A., Burnett, R. M., Darling, M. E., Lequesne, M. E., Mayhew, S. G., and Ludwig, M. L.: 1972, *Proc. Natl. Acad. Sci. (U.S.A.)* **69**, 3189–3191.

Anfinsen, C. B. and Scheraga, H. A.: 1975, *Adv. Protein Chemistry* **29**, 205–300.

Bachmayer, H., Benson, A. M., Yasunobu, K. T., Garrard, W. T., and Whiteley, H. R.: 1968a, *Biochemistry* **7**, 986–996.

Bachmayer, H., Mayhew, S., Peel, J., and Yasunobu, K. T.: 1968b, *J. Biol. Chem.* **243**, 1022–1030.

Baltscheffsky, H.: 1974a, in *Dynamics of Energy Transducing Membranes* (Ernster, Estabrook, and Slater, eds.), Elsevier, Amsterdam, p. 21.

Baltscheffsky, H.: 1974b, *Origins of Life* **5**, 387–395.

Benson, A., Tomada, K., Chang, J., Matsueda, G., Lode, E. T., Coon, M. J., and Yasunobu, K. T.: 1971, *Biochem. Biophys. Res. Comm.* **42**, 640–646.

Birshtein, T. M., Skvortsov, A. M., and Alexanyan, V. I.: 1976, *Biopolymers* **15**, 1061–1080.

Brack, A. and Orgel, L.E.: 1975, *Nature Lond.* **256**, 383–387.

Bruschi, M.: 1976a, *Biochem. Biophys. Res. Comm.* **70**, 615–621.

Bruschi, M.: 1976b, *Biochem. Biophys. Acta* **434**, 4–17.

Burgess, A. W., Ponnuswamy, P. K., and Scheraga, H. A.: 1974, *Isr. J. Chem.* **12**, 239–286.

Carter, C. W. and Kraut, J.: 1974, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 283–287.

Carter, C. W., Kraut, J., Freer, S. T., Xuon, N.-H., Alden, R. A., and Bartsch, R. G. 1974, *J. Biol. Chem.* **249**, 4212–4225.

Chothia, C.: 1976, *J. Mol. Biol.* **105**, 1–14.

Chou, P. Y. and Fasman, G. D.: 1974, *Biochemistry* **13**, 222–245.

Fox, J. L.: 1976, in *Flavins and Flavoproteins* (ed. T. P. Singer), Elsevier, Amsterdam, pp. 439–446.

Gabel, D., Rasse, D., and Scheraga, H. A.: 1976, *Int. J. Peptide Protein Res.* **8**, 237–252.

Hall, D. O., Cammack, R., and Rao, K. K.: 1974, *Origins of Life* **5**, 363–386

Herriott, J. R., Sieker, L. C., and Jensen, L. H.: 1970, *J. Mol. Biol.* **50**, 391–406.

Hong, J.-S. and Rabinowitz, J. C.: 1967, *Biochem. Biophys. Res. Comm.* **29**, 246–252.

Hong, J.-S. and Rabinowitz, J. C.: 1970, *J. Biol. Chem.* **245**, 4995–5000.

Karplus, M. and Weaver, D. L.: 1976, *Nature Lond.* **260**, 404–406.

Kobayashi, K. and Fox, J. L.: 1977, 'Abstract for the International Symposium on Evolution of Protein Molecules', Kobe, Japan, April 1977.

Lenstra, J. A., Hofsteenge, J., and Beintema, J. J.: 1977, *J. Mol. Biol.* **109**, 185–193.

Lewis, P. N., Momany, F. A., and Scheraga, H. A.: 1973, *Isr. J. Chem.* **11**, 121–152.

McCarthy, K. F.: 1972, Ph.D. Thesis, George Washington University.

Malkin, R. and Rabinowitz, J. C.: 1966, *Biochem. Biophys. Res. Commun.* **23**, 822–827.

Orgel, L. E.: 1972, *Isr. J. Chem.* **10**, 287–292.

Orgel, L. E.: 1977, *J. Theor. Biol.* **67**, 773.

Orme-Johnson, W. H.: 1972, *Biochem. Soc. Trans.* **1**, 30–31.

Perutz, M. F. and Raidt, H.: 1975, *Nature* **255**, 256–259.

Richardson, J. S.: 1977, *Nature Lond.* **268**, 495–500.

Robson, B. and Pain, R. H.: 1976, *Biochem. J.* **155**, 331–344.

Rose, G. D., Winters, R. H., and Wetlaufer, D. B.: 1976, *FEBS Letters* **63**, 10–16.

Sternberg, M. J. E. and Thornton, J. M.: 1977, *J. Mol. Biol.* **110**, 285–296.

Tanaka, M., Haniu, M., Matsueda, G., Yasunobu, K. T., Himes, R. H., Akagi, J. M., Barnes, E. M., and Devanathan, T.: 1971, *J. Biol. Chem.* **246**, 3953–3960.

Tanaka, M., Haniu, M., Yasunobu, K. T., Jones, J. B., and Stadtman, T. C.: 1974, *Biochemistry* **13**, 5284–5289.

Thompson, C. L., Johnson, C. E., Dickson, D. P. E., Cammack, R., Hall, D.O., Weser, U., and Rao, K. K. 1974, *Biochem. J.* **139**, 97–103.

Vogel, H., Bruschi, M., and LeGall, J.: 1977, *J. Mol. Evol.*, **9**, 111–119.

Ycas, M.: 1976, *J. Mol. Evol.* 7, 215–244.